# Design and Evaluation of In-Context Retrieval Exercises for Informational Videos

by

**Rimika Chaudhury**

B.Tech.,West Bengal University of Technology, 2013

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Computing Science
Faculty of Applied Science


**SIMON FRASER UNIVERSITY**
**Fall 2019**

# Approval

| | |
|---|---|
| **Name:** | **Rimika Chaudhury** |
| **Degree:** | **Master of Science (Computing Science)** |
| **Title:** | **Design and Evaluation of In-Context Retrieval Exercises for Informational Videos** |
| **Examining Committee:** | **Chair:** Diana Cukierman<br>Professor |
| | **Parmit Chilana**<br>Senior Supervisor<br>Professor |
| | **Sheelagh Carpendale**<br>Supervisor<br>Professor |
| | **Jiannan Wang**<br>Internal Examiner<br>Professor<br>School of Computing Science |
| **Date Defended:** | **25 November, 2019** |

# Statement of Contributions

This thesis includes ideas and content (including figures and tables) that have been accepted and published, in peer-reviewed publication. The conference paper from which I have adapted content is the following, which I co-authored:

Rimika Chaudhury and Parmit K. Chilana, 2019. How Learners Engage with In-Context Retrieval Exercises in Online Informational Videos. *Proceedings of the 2019 ACM Learning @ Scale* (L@S'19). ACM, Chicago, Illinois, USA.

# Abstract

Learners increasingly refer to online videos for learning new technical concepts, but often overlook or forget key details. We investigated how retrieval practice, a learning strategy commonly used in education, could be designed to reinforce key concepts in online videos. We began with a formative study to understand users' perceptions of cued and free-recall retrieval techniques. We next developed REMIND, a new in-context flashcard-based technique that provides expert-curated retrieval exercises in the context of a video's playback. We evaluated this technique with 14 learners and investigated how learners engage with flashcards that are prompted automatically at predefined intervals or flashcards that appear on-demand. Our results overall showed that learners perceived automatically prompted flashcards to be more effortless and made the learners feel more confident about grasping key concepts in the video. However, learners found that on-demand flashcards gave them more control over their learning and allowed them to personalize their review of content. Building upon findings from the design and evaluation of REMIND, we developed HYBREID to explore the design space for hybrid techniques of retrieval exercises that include the favourable aspects of automatic and on-demand interactions of REMIND. Our evaluation of this initial hybrid technique provides further implications for designing hybrid techniques of retrieval exercises. We discuss the potential for hybrid retrieval techniques where automatic exercises are combined with on-demand interactions for helping learners gain control over their study, and community support is leveraged for curating content.

**Keywords:** retrieval practice; video based learning; in-context exercises

# Dedication

My dear niece,
Anoushka,

Here I am again, with a time capsule, when I should really be dedicating fairy tales to you. But this is my gift to you: the lessons I have learned - for your journey, when you need it. To cut a long story short, here they are:
Chapter 1 gives you the bigger picture. Funnily, this falls in place only in the end.
Chapter 2 tells you to seek inspiration from the world, for it is full of wonder.
Chapter 3 is about keeping an open mind, asking questions, being curious.
Chapter 4 tells you to pursue your ideas, sincerely. And, to be a good listener.
Chapter 5 is about finding the balance.
Chapter 6 reminds you to reflect on your experience, and share it with the world.
Chapter 7 tells you to give it a closure.

The part that is not in here is to remember to have fun along the way!

Your doting aunt,

Rimika

# Acknowledgements

My parents, for instilling in me the enthusiasm to keep learning. I see the wonder in things because of you. My sister, for simply being the power-house of my life - and being that voice telling me "You can do it!" when I am not so sure.

I extend my sincere gratitude to the two most empowering women I know: my supervisor, Dr. Parmit Chilana, for placing faith in me and giving me the space to evolve into a researcher. Working with you during the last two years has been nothing short of life-transforming. And, Dr. Sheelagh Carpendale - for inspiring me to keep going back to my work with renewed energy - and filling up meetings with life-lessons-for-researchers and fun!

I am grateful to my lab mates for being a strong support (and feedback!) system - Kimia, Laton, April, Mahya, Megha and the others who joined the family later - you make the place come alive and it has been fantastic exchanging ideas with you on various occasions!

Last but not least, my friends, who have graciously endured my research related rants and helped me stay on track. Rai, you deserve a special mention, for making this thesis "roll off the tongue".

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

When trying to look up unfamiliar concepts or develop new skills, people are increasingly turning to multimedia sources and references online. Among these resources, online informational videos have become one of the most popular avenues for informal learning [11]. In fact, many learners who sign up for learning platforms, such as Massive Open Online Courses (MOOCs), often do so only to access the relevant videos [5] and therefore may not necessarily adhere to the learning pathways designed to help them learn better progressively.

While it is convenient to access informational content through online videos, many learners often find these videos to be complex and overwhelming [32, 45], making it difficult to retain the key concepts. Furthermore, informational videos about specific domains, such as information technology, can contain a lot of jargon and new concepts that learners might overlook or forget by the end of the video [10, 40]. In some cases, learners end up watching the same video several times and consulting other resources to reinforce the content they just viewed. Many of the lectures, demonstrations, and other instructional materials that learners access online rarely offer any formal structure for reviewing key concepts. Any support for reviewing content is presently offered in MOOCs such as on EdX [3] and Coursera [2], for the purpose of summative assessment to measure learning [4] and are not particularly suited for learning key concepts from individual videos.

In this thesis, we use a human-computer interaction approach to help learners better engage with informational videos and help them retain key concepts. In particular, we consider how we may offer help to the rapidly growing population of

informal learners who may not sign up for formal online courses [32], but are still interested in getting the most out of the videos that they do watch for developing specific skills.

We explored the design of interactive techniques to help learners retain key concepts in online informational videos by applying the high utility learning strategy of retrieval practice [20, 36]. The core idea of the retrieval practice strategy is that learners explicitly recall details of the content that they have previously viewed. This strategy has been shown to be effective in many scholastic learning situations [33, 35] where students are asked to reinforce their learning through intermittent quizzes, assessments, and other techniques. The other common usage of retrieval exercises is in areas which need memorization, such as language-learning [14, 19]. More recently, retrieval exercises have also been used in tasks that not only improve recall through rote-memory, but also improve a learner's ability to infer [44]. Still, little is known about how retrieval practice could be used outside of the classroom environment in more informal learning contexts.

In this thesis, we investigate the design and evaluation of in-context retrieval exercises for online informational videos from a user-centered design perspective. We first carried out a formative study with 14 participants to investigate their perceptions of common cued and free-recall retrieval exercise techniques (writing summaries, concept mapping, flashcards, multiple choice questions) in the context of watching videos related to information technology. Based on insights from this study, we designed REMIND (Retrieval Exercises for InforMaIoNal ViDeos), a novel in-context flashcard-based retrieval exercise technique that can be embedded in a video. We evaluated two versions of this technique through another user study with a different set of 14 participants and assessed learners' perceptions of flashcards that were either prompted automatically at predefined intervals in the video or accessed on-demand. Our findings revealed that learners preferred the automatic prompts over on-demand access as the automatic approach made the learners feel more confident and required less effort in determining which key concepts were worthy of review. However, many learners did find on-demand access to be useful as it gave them more control over their learning and allowed them to personalize their practice by focusing on concepts that required further reinforcement.

This key finding led us to explore the hybrid spectrum for designing retrieval exercises within video contexts such that learners requiring scaffolding may receive

necessary support automatically, without compromising their ability to have control over their study sessions. We retained the favorable aspects from the automatic and on-demand versions of REMIND and explored some hybrid techniques of practicing retrieval exercises in a second technique called HYBREID (HYBrid Retrieval Exercises for Informational ViDeos). Here we explored two combinations of automatic prompting (short pause and reminders) with granular access to exercises, and compared them with a version with the same visual cues but without any prompting. We ran a study with 5 participants to gather some preliminary insights on hybrid techniques of practice. Our key findings suggest that learners like to engage with exercises when they are visible and easy to access, and like to use them to guide their study. They also strengthen our earlier findings on the benefits of automatic prompting in further engaging learners.

Our main contribution in this thesis is in providing empirical insights into how learners perceive and engage with different retrieval exercises, when watching technical videos and highlighting the strengths and weaknesses of offering these exercises automatically or on-demand during playback. Given the recent advances in automatic assessments and AI-based learning approaches[18, 49], our findings provide useful insights into what users actually find useful and also highlight the importance of giving learners some control over their learning. We discuss several implications of our findings, including the need to explore more hybrid automatic and on-demand personalized-context retrieval exercises for informational videos.

This thesis is structured as follows: Chapter 2 outlines the related work on the theory of retrieval practice, implementations of this learning strategy in various contexts, as well as advances in video-based learning environments that incorporates types of learning strategies - including retrieval practice techniques. We discuss the elements which have inspired our prototypes, and how our work differs from them. Chapter 3 explains the formative study that laid the foundations for the design of in-context automated retrieval exercise prototypes. Chapter 4 provides details on the design and evaluation of the first prototype REMIND, and how it addresses the needs discovered in the formative study, and also sheds light on areas for further exploration. In chapter 5, based on the implications in the previous chapter, we further explore the design space for in-context retrieval exercises with some hybrid techniques using our second prototype - HYBREID. Finally, in chapter 6 and 7 we reflect on the advantages and

disadvantages of the approaches in the aforementioned prototypes, and discuss other implications from this research and potential future work.

# Chapter 2

# Related Work

Our work builds upon research in learning sciences and human-computer interaction (HCI) with a particular focus on the theory and applications of retrieval practice. Our work also is also inspired by a wide range of innovations in the video based learning contexts.

## 2.1 The Theory of Retrieval Practice

Retrieval practice is identified as one of the high-utility learning techniques that are often used in educational contexts [36, 37]. The benefit of retrieval practice is that it makes few demands from the learner and promises high learning gains. Cognitive science explains retrieval practice as a way to promote learning by using low-stake "tests" or prompts to retrieve recently viewed concepts or facts from memory, As a result, it is often referred to as the "testing effect" [37]. The usual usage of the term "testing" implies that a student is assessed using standardized tests but Roediger et al. [37] review the testing that occurs while a learner is in the act of studying and thus uses the exercise as a form of self-testing. As opposed to using testing as a way to determine what the learner knows, this study showcases the other effect of testing - that it changes what the learner knows and contributes to better retention of the material on which they are tested. In the context of our work, where learners often tend to restudy the material [20] by rewatching the video over and over again, we use the idea of self-testing instead, for improving retention of the key ideas from the video.

A study by Schmidmaier et al. [38] that investigated the importance of retesting strategies in the learning of factual knowledge in medical education found that practicing the facts through repetition using electronic flashcards improved the short-term retention of the material. The improvements after repetitive testing (as opposed to repetitive studying) were observed within a span of one week. Another study by Roediger et al. [35], also examines the effect of self-testing versus re-studying, on long-term retention of the material. From these studies, in general, retrieval practice is shown to be superior when compared to repeated restudying, and impacts the learning by improving retention of the tested material. The benefits of retrieval practice also extend towards enabling learners to make inferences. In a study by Smith et al. [44] it was found that when learners practiced retrieval, they showed improved retention on both recall-based as well as inference-based questions.

However, since these studies used different spans of time to assess retention, the context of "short-term" and "long-term" become relative. Other studies have further explored the idea of spaced retrieval practice and found that exercises that are spaced out over a span of time are more effective than the ones that are close together [9]. Most of these studies have been carried out in formal education settings and have focused on testing wide gaps of intervals, ranging from days and weeks over months. Still, given the learning gains possible through retrieval exercises, we explore this idea to see if learners would find such exercises useful in the context of learning from a single informational video. In our work, although we employ strategies for encouraging repeated practice at different intervals of time, our study conditions span only a single study session.

## 2.2 Innovations in Implementing Retrieval Exercises

Retrieval exercises can broadly be categorized into: 1) free recall, where the learners do not receive any hints while attempting the retrieval, and 2) cued recall, where some form of support is provided to help the learners recall better [44]. Some popular formats include braindumps, quizzes and flashcards [20, 25, 28].Several innovations in learning have explored how retrieval exercises can be designed in different formats (e.g., free and cued recall). Smith et al. [43] investigate how presenting retrieval exercises in different formats may affect learning and compare between short-answer style

which is common in scholastic settings, multiple-choice questions which is common in MOOCs, and also a hybrid variety that combines the two. They found that the format of the exercises do not impact learning outcomes. Research suggests that the act of attempting to recall itself provides benefits regardless of the success of the exercise [7]. Retrieval exercises have been shown to be beneficial to learning, even in the lack of feedback [44]. However, some works have found that the effect of retrieval practice could be attenuated depending on the material and its style of presentation [15]. Although the assistance provided during cued recall does not increase the benefits from the retrieval act itself, it does allow the learner a chance to correct their understanding. Given the benefits of retrieval exercise, irrespective of the format, in our work, we explore which kinds of retrieval techniques are perceived to be engaging in video watching contexts. We designed our formative study to initially assess learners' perceptions of retrieval exercises in multiple formats.

Learners typically also have to navigate through a great deal of content within a limited time and therefore the time spent on practicing is also an important consideration. Pyc et al.[34] have particularly looked at the scheduling aspect of retrieval practice, where they compare conventional scheduling, which is consistent with most previous research, with a dropout schedule. The conventional schedule consists of an initial study phase followed by three practice phases, whereas the dropout schedule makes the study more selective by discontinuing the practice of learned items. In our work, we exploit the advantage of a drop-out schedule in our initial prototype (REMIND) by allowing the learners to refine their practice towards items they identify as pain-points. A study by Karpicke [20] highlights that learners may not be able to make good choices about practicing retrieval once they judge that they have learnt the item and that may lead to poor retention eventually. In our work, we explore two kinds of designs for retrieval practice within informational video context, through our prototypes: In our prototype REMIND, the learners are cycled through retrieval items with limited choice for selecting items, and in HYBREID, they have more freedom for selections.

Investigations have also tried to identify how suitable retrieval exercises can be adapted and integrated into various learning contexts [6, 10, 18, 20, 45]. For example, a study by Becker et al. [6] looks at how meaningful fill-in-the-gap styled retrieval exercises may be automatically generated from text material. In another study by Huang et al [18], automatic multiple-choice-quiz generation is explored for TED talks

video clips. Davis et al.[10] designed and evaluated an automated adaptive system for delivering retrieval practice in the form of multiple-choice questions within a MOOC on a weekly basis and in this study they uncovered several aspects about designing retrieval practice for digital environments for better learner experience as well as knowledge retention.

A limitation of the existing works on utilizing or exploring retrieval exercises is that they have been done in formal scholastic settings, mostly involving students enrolled in courses or a more mixed audience for MOOCs [1, 10] or in language learning settings [14, 40, 50]. In contrast, we investigate learners' perceptions and explore designs for retrieval exercises in non-scholastic situations, such as with the use of informational videos, where learners are more self-directed and have specific motivations for learning particular concepts. In a more formal context, the learner commits to following a course pathway that usually build towards gaining a certificate. In this case the forms of assessments that are offered in the course are used towards summing up the performance of the learner.However, a learner may also sign up informally, simply out of curiosity, or general interest [24] and it has been found in a study conducted by Shrader et al. [41], that this is the case for the majority of the enrolled members. In this case, the assessments become a means towards enhancing and supporting their learning, besides being means of assessing their knowledge. In the latter case, where the learners are not bound by the formal structures any more, self assessments are found to be an important technique to support their learning [46]. In our work, we explore this specific case of offering self-assessments by looking at various retrieval practice techniques in video watching contexts to shed some light on how this experience could potentially be designed.

## 2.3 Innovations in Video Based Learning

Recent innovations in HCI are also tackling the space of improving learning with online videos through various interactive strategies. Some examples include the integration of interactive components within the context of a video, including comments, threads, assessments[30, 47], prompts with reflective questions [26, 48], and interactive notetaking [12]. For example, a study by Dimitrova et al. [12] explores a novel design for engaging the learner in active video watching through the means of interactive note-taking. These interactions are then used to automatically identify appropriate

portions of the video to provide nudges, in order to actively engage the learner with the video. In our work we borrow the idea of nudges in the form of reminders and deliberate pauses to prompt the users to engage in retrieval exercises. A study by Kim et al. [21] looks at rich interactive medium for involving both the teacher and student in the learning activity. One work by Kovacs [26] which is close to what we try to achieve with one of prototypes, offers a question based format for navigating a video and helps learners keep track of their question-answering history. Although in our work video navigation was not the primary goal, we have used visual components to assist the learner in making a decision about engaging with an exercise.

Prompting an exercise within the context of a video, while the learner is in the process of watching is also a theme that Liu et al. [27] explore in their research. While we share the goal of understanding learners' perceptions of in-video exercises, in our work, we also try to identify the elements of retrieval practice exercises that are useful to learners such that the exercises can be designed in a way that will promote retention of the key concepts. The lack of support in video watching interfaces to effectively learn the contents of the video has been identified by several researchers before [13, 47, 48] and solutions such as assessment integration for supporting learning has been explored. In the research conducted by Dodson et al. [13], we see efforts in the direction of designing an active video watching environment, while Seidel [39] provides us with a comprehensive insight on different ways assessments may be integrated within videos depending on the specific issues that learners may face while learning from videos.

These works suggest that integrating interactive elements to augment the video watching activity positively enhances not only learning outcomes, but also the overall learner engagement. Our work complements these existing works by studying and designing interactive designs for retrieval exercises that can be integrated within videos.

We focus on learners' perceptions and how they engage with different techniques to derive design insights rather than measuring learning outcomes resulting from the use of these techniques. Integrating interactive components within the context of video introduces the possibility of breaking the flow for the learner. Studies suggest solutions to the problem of inserting these break-points in the video in a suitable manner [39] and some innovations [12] have tried to predict suitable points. We have taken inspiration from spaced-retrieval using flashcards [25] and use expert-curated intervals, and also encourage learners to practice older items to strengthen their weak-

points [26]. Finally, some recent innovations have tried to enrich the video-based learning experience by exploring collaborative designs, including crowd-sourcing [27], teacher-student oriented feed-back [21] as well as peer-to-peer [12] based techniques. Although our current investigation focuses on assessing the perceptions of individual learners, we offer insights for future work to explore collaborative designs for designing curated retrieval exercises at scale.

# Chapter 3

# How Learners Perceive Retrieval Exercises: A Formative Study

As discussed above, retrieval exercises have been shown to be beneficial in improving learning outcomes in formal education settings where practice can be controlled and enforced. However, there are few insights into how such retrieval exercise techniques should be designed in more informal learning contexts, such as watching informational videos.

The goal of our formative study was to compare between different styles of retrieval exercises and assess how well learners engage with these techniques when watching informational videos. In particular, we focused on two main categories [44] of retrieval exercises: free recall and cued recall. During free recall, the learner is tasked to reproduce the material entirely on her own (e.g., writing summaries, concept mapping). During cued recall, the learner receives a cue that serves as an aid for memory retrieval (e.g., multiple choice questions (MCQ), flashcards).

Our overarching research question was: *What are learners' perception of cued and free-recall retrieval exercises when watching informational videos?*

We designed a lab-based observational study with 14 learners and also collected user feedback through survey responses and follow-up interviews, as described below.

## 3.1 Study Design and Procedure

During the study, we asked participants to watch short video lectures on popular technical concepts (Table 3.1) selected from *Lynda.com*. The content of each video mostly targeted learners on-the-job and we made efforts to maintain consistency by selecting videos that were similar to each other in terms of duration, style of presentation and in terms of number of concepts discussed. Following each video segment there was an associated learning activity (5 minutes), using one of the aforementioned retrieval techniques: free recall (writing summaries, concept mapping) and cued recall (MCQ and Flashcards). The study was conducted in a lab setting using the think-aloud protocol where participants were encouraged to share their thoughts throughout the study.

The retrieval exercises were generated using *Microsoft OneNote* as it offers various features for writing text, drawing, and adding forms. The free recall tasks required some typing interaction (see Fig. 3.1 and Fig. 3.3). The cued recall tasks only required click interactions (see Fig. 3.4 and Fig. 3.5. The participants were not provided with any review content or feedback (correctness of responses) after the free recall tasks. However, for the cued recall condition, the tasks were created using a list for MCQ and some text based flashcards where the cues themselves acted as review content. Optionally, participants could check the correctness of their response for the cued recall exercises if they wanted.

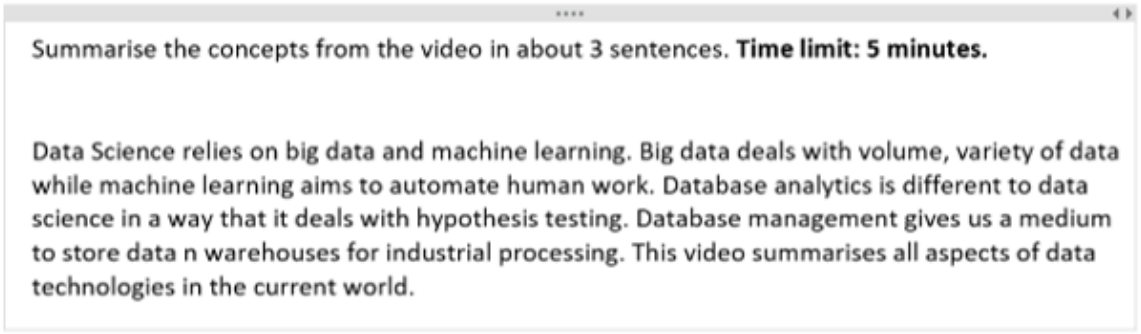


Figure 3.1: Response recorded by P13 for summary-writing exercise.

Following each condition, we asked participants to complete a survey related to the learning activity and rate their experience on a 5-point Likert scale and we also

conducted a final short interview where we probed them about their experience and perceptions. The survey and interview questions focused on selected proxy engagement measures [31, 42], such as whether learners found these exercises to be helpful, enjoyable, distracting and to what extent the exercises encouraged reflection, offered a sense of control, and boosted learner confidence (Fig. 3.2).

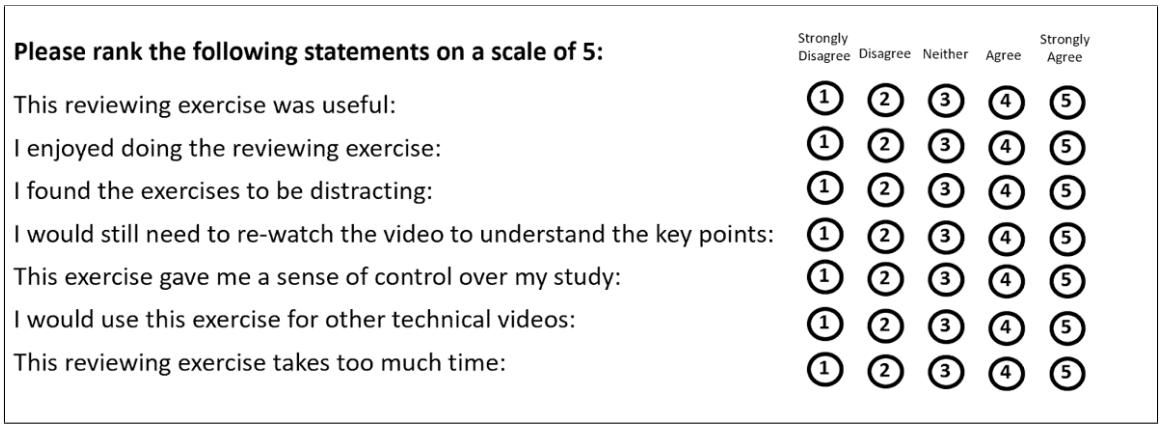


Figure 3.2: The survey questions answered by the participants for the formative study.

## 3.2 Participants

There were 14 participants (8 male, 6 female) and all were university students within an age range of 19-31 and came from Computer Science or Information Technology backgrounds. All of the participants had experience in consulting technical videos outside the classroom. They received $15 Amazon Gift Cards for an hour-long study. The retrieval techniques and the videos were paired as shown in Table 3.1. We followed the Latin Square arrangement to expose the participants to the treatments to counter any order effects. Conducting one-on-one user study allowed us to draw out the deeper reasons behind the pros and cons of the various retrieval techniques from the participants' perspective.

## 3.3 Analysis

Our evaluation used a mixed-methods approach to collect data in multiple ways: questionnaires, observations, think-aloud and interviews. We first analyzed the eval-

| | | **Test Conditions** |
|---|---|---|
| *Video Chapter* | *Technique* | *Condition* |
| IoT Foundations | Flashcards | C1 |
| Data Science | Summaries | C2 |
| Programming Foundations | MCQ | C3 |
| Cloud Architecture | Concept-Map | C4 |

Table 3.1: Four retrieval exercise techniques were mapped to four videos.

uation questionnaires to compare results between the two conditions. We also coded the interview transcripts and observational notes to discover recurring themes using inductive analysis approach [8]. Next we present the key findings about the learners' perspectives on retrieval exercises for informational videos.

### Retrieval Exercises Promote Thinking while Watching Videos

Overall, participants indicated that *"all of the activities [exercises] were engaging"* (P02) as these exercises gave them a chance to pause and think more deeply about the concepts shown in the video. Participants noted that the exercises made the ideas concrete or made them realize immediately what they could not recall: *"without the activities [exercises], I would simply end up re-watching the videos again and again"* (P02).

Although, most of the participants (11/14) indicated that the exercises were overall not distracting, some participants did mention that free recall exercises, such as summaries, *"should not appear in the middle [of the video] when I am already thinking of something...preferable to have something simple there.."* (P03). Many participants (5/14) also indicated that they did not enjoy the free recall exercise of concept mapping.

### Less Effort Made Cued Recall More Enjoyable

Participants in general were more positive about cued recall exercises compared to free recall. The overall consensus was that cued recall required less cognitive effort to remember key details: *"It is hard to recall when a topic is new"* (P13) and participants also felt having cues as a *"confirmation [of their understanding] would be useful"*

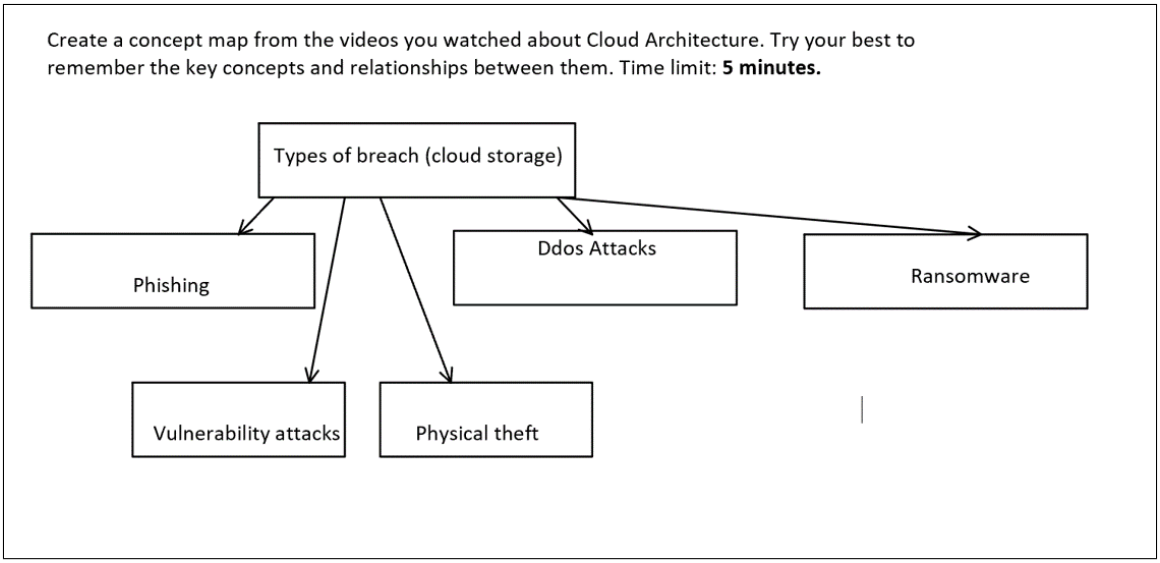


Figure 3.3: Response recorded by P13 for concept-mapping.

(P06). Although participants felt that free recall would be more beneficial for their learning in the long run, free recall exercises in the context of watching videos seemed like *"too much work"* (P12) and most participants did not feel satisfied with their recall. In contrast, the presence of hints in the cued condition and the option to check the correctness of their responses served as useful feedback for the participants:

> *"...Writing summary was my least favorite... least useful. No feedback does not help me."* (P02)

### Learners Like Control with Flashcards

With cued recall emerging as the favored retrieval exercise technique, we looked closer to compare MCQ and flashcard styles. We found that 9 out of 14 participants reported enjoying the MCQ exercises (Fig 3.4)because the *"MCQ [format] was familiar"* (P10) and *"there are the other options that give you more hints"* (P13). However, many participants rated MCQs as being more complicated than flashcards due to the level of detail in the question and sometimes confusing or tricky options. They preferred a simpler representation of the content. Of the participants who agreed that cued retrieval techniques encouraged reflection of key concepts, 2 participants mentioned

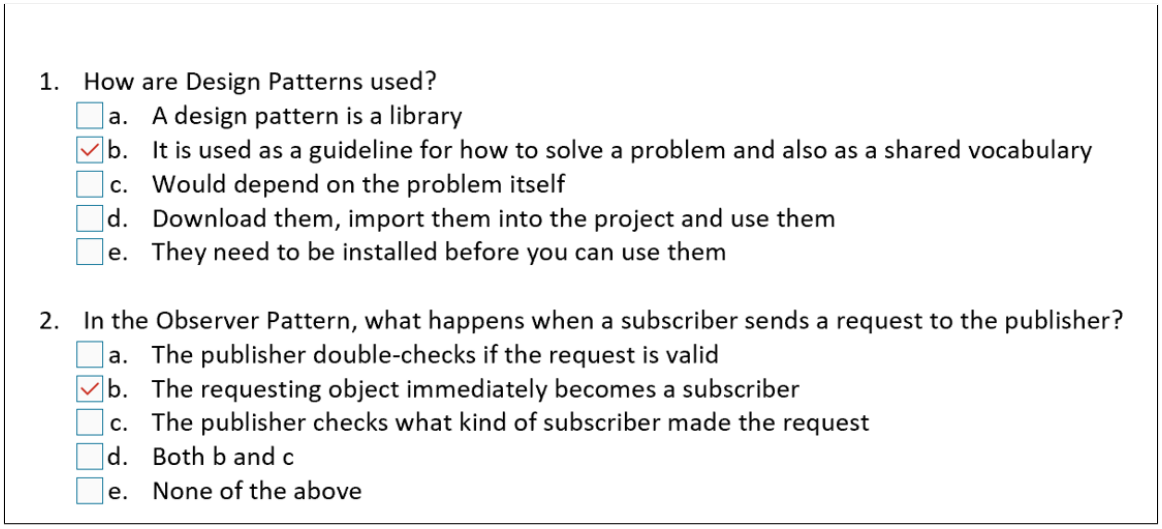


Figure 3.4: Response recorded by P13 for Multiple Choice Questions (MCQ).

that MCQ did not give them "*much control*" and that these exercises did not always provide *"enough practice"*.

In contrast, participants perceived flashcards to be more favorable (Fig 3.5). In fact, 10 out of 14 participants agreed to some extent that they felt more *"in control"* of their study as they could review cards several times before testing themselves. They found the material on the flashcards to be less distracting and overall found them to be useful. For example, one participant explained: *"[flashcards] are useful in case you forget some material and self-test helps confirm if you get it"* (P05). Participants also mentioned that they would prefer flashcards when they have a *"[low] attention span or... time is a constraint"* (P01). Participants perceived flashcards to provide a comprehensive review and thus were more useful in helping them recognize their pain-points, as opposed to MCQ which offered little benefit for repeated practice.

## 3.4 Design Implications for Retrieval Exercises

Based on the formative study insights, we identified the following implications for designing engaging retrieval exercises for informational videos:

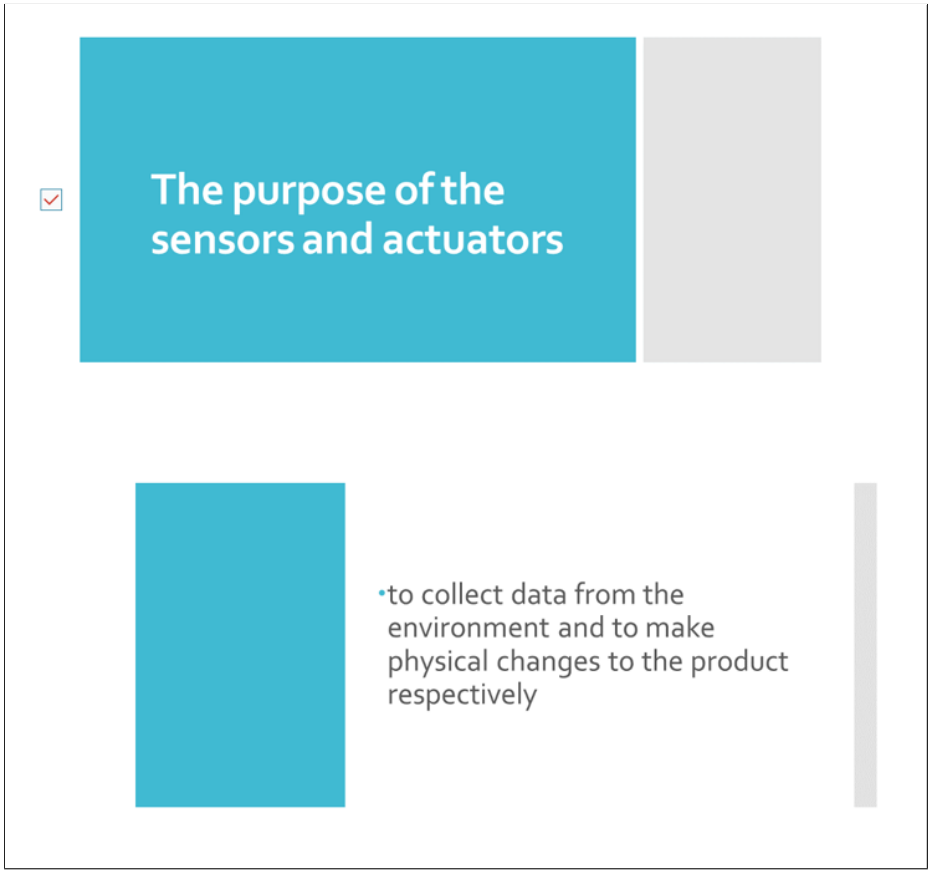


Figure 3.5: For the flashcard exercises, clicking on the first card (cue) revealed the second card (explanation). Learnt cards could be marked with a check.

### D1. Minimize recall effort

A key takeaway from the formative study was that learners perceived cued recall, with flashcards in particular, to be more enjoyable and less effortful than free-recall. Cued recall exercises should be designed for informational videos to be easy to perform, require minimum interaction, and require minimal knowledge construction from scratch.

### D2. Provide curated exercises

We found that content that is curated by experts was a trusted source of feedback and actively sought by learners. Hence, retrieval exercises in informational videos that provide curated feedback would be more useful for learners.

#### D3. Allow Learners to have more control

While learners find curated exercises useful, they also want some flexibility in selecting the content and the frequency of their review based on their evolving understanding. Retrieval exercises in informational videos should offer learners control over their pace of learning and opportunities to reflect on what they do or do not understand.

## 3.5 Summary

In summary, we learned that learning new content itself is challenging and therefore curated retrieval exercises within videos help learners maintain focus by providing moments for reflection. We also learned that exercises that require more effort in terms of implementation, are less desirable in video watching contexts. Cued recall exercises provide challenges of adequate retrieval difficulty while also offering feedback, which is an aspect of exercises that learners find desirable. Finally, besides having access to curated exercises, learners also prefer to exercise control over their study session. Therefore, in the next chapter, we look at ways of designing interactions to allow learners to control their retrieval practice sessions.

# Chapter 4

# Design and Analysis of In-Context Flashcards in REMIND

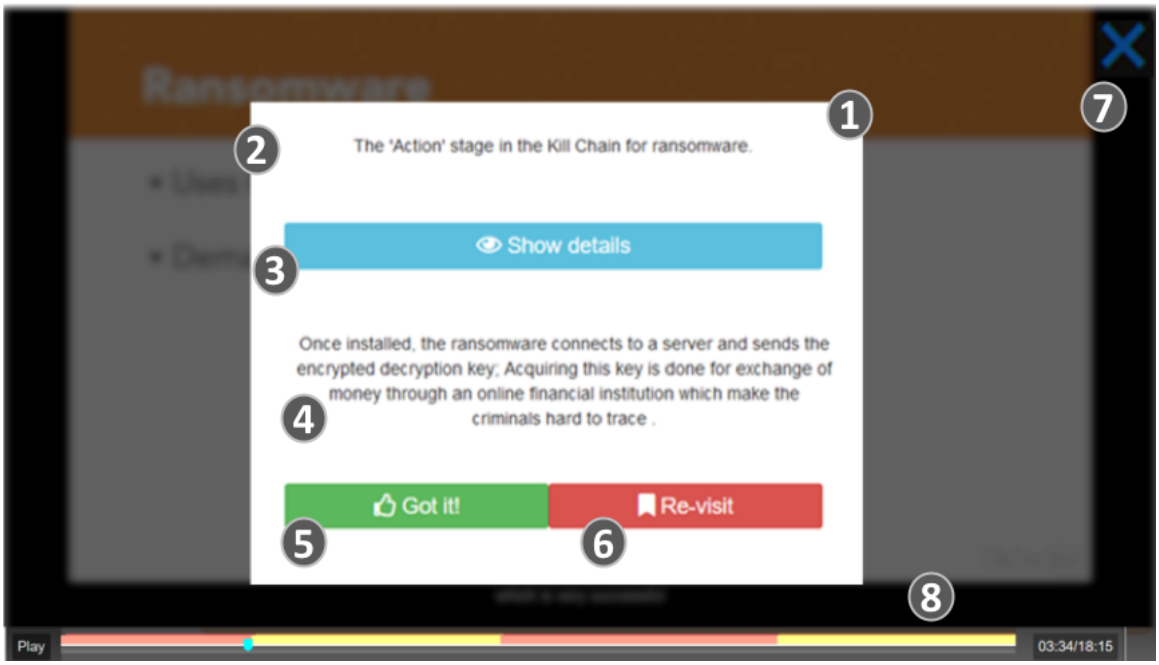


Figure 4.1: The interface of REMIND illustrating a curated timestamped exercise.

Informed by the design implications in the previous chapter, we designed a novel web-based retrieval exercise technique consisting of interactive curated flashcards that can be retrieved in-context of video playback (Fig. 4.1). In this chapter, we present the design of REMIND (Retrieval Exercises for InforMaIoNal ViDeos). We also describe our study procedure for evaluating the prototype, and discuss our key findings.

## 4.1 Interactive Curated Flashcards

The key idea motivating our flashcard retrieval exercise technique was that a learner should be able to interactively reflect on key concepts at different intervals within the video. In line with D2, we offered flashcards that were curated by subject matter experts.

Each flashcard consists of two statements: Statement 1 (S1) presents a concept discussed in the video (Fig. 4.1.2) and provides the learner a general statement about the concept as a cue *(D1)*. Statement 2 (S2) (Fig. 4.1.4) reveals more minute details relating to the concept in S1 *(D2)*. By default, S2 remains hidden and is only revealed when the learner clicks on the "Show details" (Fig. 4.1.3) button. S2 may contain specific keywords, jargon, terminologies or precise examples and instances of the concept in S1. The choice of viewing S2 for details allows the learner to control the depth of the retrieval exercise *(D3)* and provides additional confirmation if desired by the learner.

If a learner is satisfied by her understanding of the the concept on the flashcard, she can select the "Got it!" button (Fig. 4.1.5), and move on to the next flashcard. Otherwise, she can bookmark the concept by clicking on "Re-visit" (Fig. 4.16) and the card will be re-retrieved in the next set of flashcards. There is also a dismiss button (Fig. 4.1.available to allow learners to dismiss the exercises at any stage of completion and continue watching the video. In line with D3, the learner has control in assessing her comprehension, reflecting on her weaknesses, and doing no reviews or additional reviews when necessary. Lastly, a comprehensive practice at the end of the video is also available to allow the learner to run through all of the flashcards again for one final self-assessment.

## 4.2 In-Context Access to Relevant Flashcards

With our interactive flashcard technique, a new flashcard becomes available after a concept has been discussed in the video. A time-stamp of the moment that particular concept was covered in the video gets attached to the flashcard. The segmented seek-bar, in Fig. 4.1.8, shows different segments in the video where flashcards would be available. For example, in Fig. 4.1, the learner has completed watching a quarter of

the video and has arrived at the end of the first segment. At this point, the learner gets a set of flashcards that have a smaller time-stamp than the the current moment in the video. At the next stopping point in the video, a new set of flashcards become available for practice along with a set of cards from the previous segment(s) that the learner had bookmarked (if any). This design pertains to the goal of providing curated exercises *(D2)* with timely availability for engaging the learner *(D1)*.

In designing our flashcard-based in-context retrieval technique, a key question we faced was when and how often would learners actually want to interact with retrieval exercises when watching informational videos? To investigate this, we explored two designs for our flashcard retrieval exercise technique: 1) *Automatic Prompts*: flashcards that appear automatically in the video at predetermined intervals decided by experts; or, 2) *On-Demand Prompts*: flashcards that can only be retrieved on-demand by learners when they feel the need to review or reflect on a concept.

## 4.3 User Evaluation of REMIND

To assess the design of our flashcard retrieval technique and to compare the two possibilities for offering the flashcards in context of informational videos, we designed a comparative observational user study (Fig. 4.2). We sought to answer two main research questions:

1. *What are learners' perceptions of the utility of interactive in-context flashcard-based retrieval exercise embedded in informational videos?*
2. *What are learners' perceptions of flashcard exercises offered automatically, at intervals prescribed by experts, versus flashcard exercises that are accessed on-demand by learners on their own?*

### 4.3.1 Study Design

We used a within-subjects design for our study, where each learner used both versions (automatic and on-demand prompts) of the flashcard-based retrieval technique embedded in two different informational videos (Table 4.1). We worked with experts

Figure 4.2: 1) The figure on the left shows 1) Automatic Prompts condition, which contains the segmented seekbar for topic transitions and automated pauses at every transition. The figure on the right shows 2) On-demand Prompts condition that shows that the segmented seekbar is absent and the mouse-pointer depicts that the learner needs to access the flashcards on pausing on their own

| | | **Test Conditions** |
|---|---|---|
| *Video Chapter* | *Technique* | *Condition* |
| Cloud Security | Automatic Prompts | C1 |
| Cyber Security | On-Demand | C2 |

Table 4.1: Automatic and on-demand flashcard techniques were mapped to two videos

to first identify key concepts in each of the videos and accordingly created a flashcard for each of these concepts (total 12 for each video). Each concept on a flashcard (Fig. 4.1) was associated with a time-stamp to facilitate the retrieval of relevant flashcards at different intervals during playback.

### Automatic Prompts Condition

In the automatic condition, a set of flashcards automatically appeared as an overlay in the video at intervals curated by experts. These intervals were natural topic breaks in the video and experts created flashcards relevant to content that would appear within 4 roughly equally-spaced intervals. The intervals were indicated on the seek-bar with alternating colors to provide a visual cue (Fig. 4.1.8). The video automatically paused at the end of these intervals and a precursory prompt with a link to access flashcard exercises was shown to the learner. The whole set of qualifying cards (based on the time-stamp) was retrieved where there were two to three flashcards in each set, relevant to the concepts discussed in the immediately preceding interval.

**On-demand Prompts Condition**

In this condition, no flashcards or intervals were shown to the learner automatically- the learner could access flashcards by explicitly pausing the video at any point. The pause would then trigger the same overlay to extend over the video and display the precursory prompt with a link to access flashcard exercises. Since each flashcard was associated with a specific time-stamp, it would only became available for review after the learner had viewed the relevant portion of the video. The visual cue using colors to indicate topic breaks was removed in this condition. Other features of the interactive flashcards described earlier, such as bookmarking and skipping, worked the same way in both conditions.

### 4.3.2 Procedure

We recruited a new set of 14 participants (7M, 7F), all between the ages of 19-40 who either had a computer science related education, or relevant technical experience. All participants were given $15 Amazon Gift Cards for their participation.

Each participant first began the study by filling out a demographic questionnaire. Next, participants interacted with the two different flashcard techniques described above. To counter order effects, all participants were randomly assigned to either the automatic or the on-demand condition. Before starting either condition, participants were given the following scenario: they were newly hired employees of a technology company which expected them to come up to speed with some new technical concepts for their upcoming project. The participants were then given a tutorial of the flashcard-based exercises and asked to use it just as they would if they were in the aforementioned scenario while watching two 18-minute long video lectures - one for each condition.

The participants were asked not to take notes and were instead encouraged to make the best use of the hints on the flashcard to recall the concept-related details. We introduced the bookmarking feature and encouraged participants to use it until they were comfortable with the concept. There was no time limit for the exercise sessions or for completing the surveys, but we maintained an upper-bound of 1.5 hours for the entire study. The videos were sourced from Lynda.com (see Table 4.1)

and were similar to each other in terms of duration, style of presentation and the number of concepts discussed.

For the automatic condition, the participants were informed about the visual cue on the video seek-bar that indicated the segments in the video and the points where a set of flashcards would appear automatically. They were free to either take a suggested review at the end of each segment or dismiss it if they felt they did not need it but were asked to take at least one exercise at any point. For the on-demand condition, the participants were asked to pause the video on their own and access the flashcards for practice whenever they felt the need for it. In this condition also, they were asked to take at least one exercise at any point.

The study was conducted in a lab setting and the Think-Aloud protocol was followed where the participants were encouraged to share their thoughts during the study.

## 4.4 Data Collection and Analysis

We collected data by directly observing participants' behavior during the study, recording the screens, collecting their feedback through survey responses and drew further insights from interviews which were recorded and later transcribed.

In this evaluation, we allowed learners to have more freedom than our formative study by allowing the participants to choose how many times they wanted to engage with the flashcard exercises. We also made observations to find out whether the learners were cognitively engaging with the exercises by recording their screen and noting their bookmarking behavior, how many times they dismissed an automatically prompted exercise or voluntarily stopped to view an exercise set, how often they saw the details in each card, and also the total time spent in engaging with exercises. Similar to our formative study, we also used selected proxy engagement measures [31] which were relevant to our study. We were mainly interested in finding out if our in-context flashcard exercises were useful, enjoyable, offered a sense of control over their study, boosted confidence about the material, seemed time consuming or distracting and promoted focused attention 4.3.

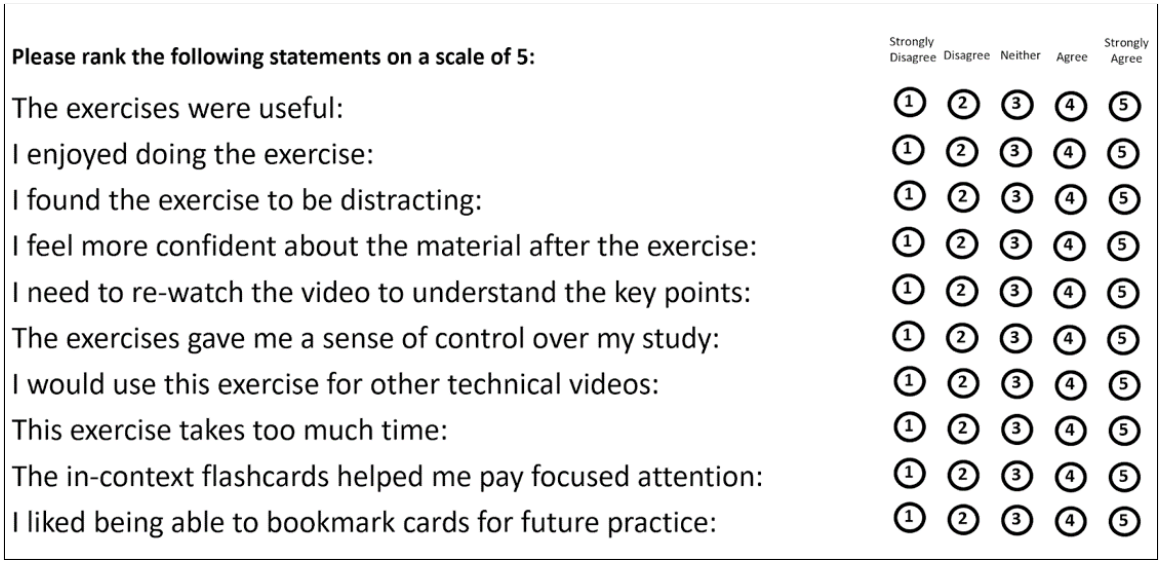


Figure 4.3: 1) Survey questions for evaluating REMIND.

### 4.4.1 Results

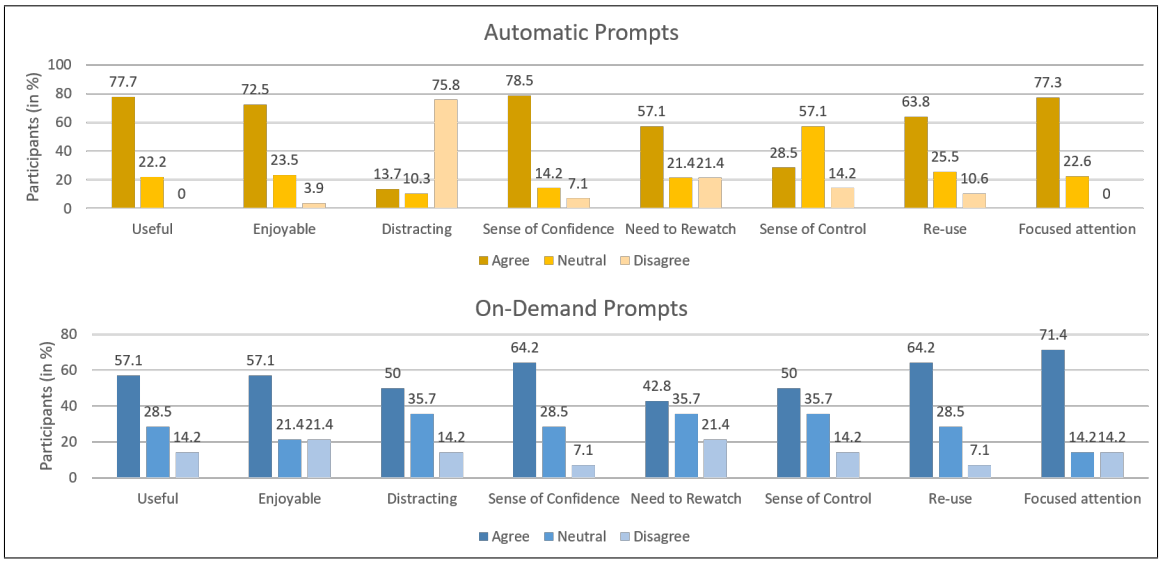


Figure 4.4: Study results for user-engagement measures for Condition 1: Automatic Prompts and Condition 2: On-Demand Prompts. Participants rated the review exercises on a 5-point Likert scale ranging from Strongly Disagree (Rating 1) to Strongly Agree (Rating 5). In the above figure, Strongly Agree and Agree responses are added together and labelled as Agree. Similarly, Strongly Disagree and Disagree are clubbed and labelled as Disagree.

We present our main findings about how flashcard-based retrieval exercises were perceived when they appeared in context of an informational video and how the automatic delivery of the flashcard exercises compared with on-demand access.

### Utility of In-Context Flashcard Exercises

We found that most participants (10/14) agreed to some extent that they found in-context flashcards to be useful. They cited many reasons, such as being able to quickly remind themselves of the key concepts before progressing on to the next sections in the video. The flashcard exercises also helped participants confirm their understanding: *"The flashcards gave me a more solid understanding of what the topic was [all] about"(P03)* Flashcards showed potential for use in conceptual or informational videos, especially those that contained new information for the learner:

> *"...watching information based [videos] where they are trying to transfer some kind of knowledge, these kinds of [flashcard] reviews help. When you are less familiar with a topic, you need to retain more and need to grasp more"(P11)*

Participants mentioned that in situations where they experienced difficulty in grasping the concept, they tended to re-watch the video several times. However, after taking the flashcard-based retrieval exercises, over half of the participants (8/14) disagreed or were neutral about feeling the need to re-watch the video. One participant explained: *"You could actually do away with the video after just one watch."* (P12). Overall, flashcards provided a useful alternative to re-watching the videos, *"If I am in a hurry"* or *"if I am looking at a conceptual material"* and *"do not want to go through the whole material"* (P12). The participants displayed a high level of engagement by staying committed to the flashcard exercises once they had decided to take them.

Many participants (8/14) made use of the bookmark feature in the automatic condition (34 bookmarks) and 6/14 participants used bookmarks in the on-demand condition (26 bookmarks), generating 60 bookmarking events in both conditions combined. This shows that the learners were not only actively engaging in the exercises, but also regulating their understanding and planning for repeated practice on their own (without being prompted by the experimenter). The participants confirmed in

the interview that they found bookmarking to be a useful feature: *"If you are just doing the final flashcard review in the end, and if you are doubtful about it, you can bookmark it, go through the whole thing and have it pop up again."(P04)*

Additionally, most participants (10/14) in either condition attempted the optional final review at the end of the video, suggesting that they were engaged with the final retrieval exercise and did not rush to finish the study.

### Accessing Flashcards Automatically vs. On-Demand

To compare both the automatic and on-demand conditions for accessing flashcards, we analyzed the data collected through direct observations, survey responses and by interviews. We focused on user interaction and behavior in both conditions as well as perceptions of effort, confidence, and control.

**User Interaction and Behavior:** In the automatic condition, we observed that participants' engagement with the exercises was consistently high across all of our key metrics (Fig. 4.4). Although participants were required to attempt only a minimum of one exercise and could skip the rest, it was encouraging to see that all 14 participants took 3 or more exercises when automatically prompted (in fact, 12/14 participants took all 4 exercises). Survey feedback indicated that only one participant found these exercises to be distracting when watching the informational video and the distraction mainly seemed to stem from the colors in the video player seek-bar (Fig. 4.1.8).

The rationale behind the on-demand access was to provide the learners with more control over their study session - allow them to pause for reflection more naturally, when they felt the need for it. We observed that most participants (10/14) proactively paused to take the retrieval exercises during the on-demand condition, suggesting that retrieval exercises were helping the participants to grasp key concepts in a progressive manner. We also observed that the participants varied greatly in how frequently or regularly they paused - mostly ranging between 2 to 5 times, but in extreme cases - never pausing in between [P3, P6, P8] or pausing up-to 12 times (P10). However, unlike the expert-curated intervals in the automatic condition, the user-selected pauses rarely aligned with the natural topic-breaks in the video. This behavior suggests that the participants were indeed relying on their natural sense of flow to determine when to stop, which may not necessarily align with topic-breaks in the video.

**Perceptions of confidence in understanding content:** In the automatic condition, most participants (11/14) agreed that they felt more confident about the content after taking the flashcard exercises. The participants felt that the retrieval exercises at pre-defined intervals provided a beneficial guidance. Qualitative insights confirmed that the automatic prompts helped the participants gain more confidence about the material and automatic prompts at regular intervals *"made the ideas stick"* (P04) as they progressed through mounting details. This also helped them gain better value out of the comprehensive exercise in the end. *"[Automatically prompted flashcards] helped me recollect things I forgot from the earlier segments."* (P02). *"I was able to have small recaps before I could take the final one.. so I was more confident about the particular sections made the comprehensive review at the end more useful"* (P04). In contrast, during the on-demand access condition, fewer participants (9/14) agreed to feeling confident about the material after taking the exercises. Our interviews revealed that participants were more uncertain in the on-demand condition as to whether they had actually stopped at an appropriate point to seek review in the video.

**Perceptions of effort:** As we noted, half of the participants (7/14) in the on-demand condition only took one or two of the four possible exercises. We found that although the participants liked having the freedom to pause at any point for a retrieval exercise, there were cases where the participants felt that having to pause on their own added to their efforts of trying to learn from the video. The on-demand access appeared to add more effort because it was difficult for learners to identify suitable places to stop: *"I have to tell myself I need to stop and take some exercise. Before watching the video I was planning to pause 3 times but later I found that some important things are going on and I cannot pause here."* (P14).

In fact, some participants were unable to identify natural topic breaks at all due to the absence of visual cues: *"The second video had more concepts... but no [topic] segments."* (P03). The automatic prompts also made the decision to stop for exercises easier: *"better if it [retrieval exercise] is forced on me"(P08)* and *"I am lazy... so I wouldn't pause on my own if I know [the flashcards] will appear automatically."* (P03).

**Perceptions of control over learning:** Despite the overall merits of the automatic condition, a key drawback noted by participants was a lower sense of control over their study. Although a learner could dismiss or postpone a prompted exercise and bookmark cards for more practice, only a few participants (4/14) agreed to some

extent, feeling in control. A key reason for this was that the participants felt that the predefined timing for the prompts did not always align well with their need for a review: *"It was not the right timing for me... too many pauses slow me down."* (P07)

The on-demand access of flashcards in the second condition made up for the shortcomings of the automatic prompts and control. Participants overall appreciated having control over their learning by being able to access exercises whenever they felt the need to reinforce the concepts they just learned: *"I prefer the [on-demand design] where I have control because I don't want to get interrupted. I like the flow."* (P04) Furthermore, with the ability to access the exercises on demand, participants said that when they *"...found [my] attention wandering...it helped [me] slow down and pay more attention."* (P07).

We also observed some instances where participants wanted even more control: *"I prefer to be able to choose what [I want] to check"* (P06). This participant preferred to see only the most immediate flashcard, instead of revisiting all cards.

Participants overall indicated that although they had a slight preference for the automatically prompted flashcards, the ideal case would be one in which they have *"best of both worlds"* (P12), meaning on-demand access was a desirable feature but best when augmented with automatic prompts.

## 4.5 Summary

The evaluation of prototype REMIND provided us with insights that learners neither lean heavily towards automated interventions nor prefer to take full control of their practice. Most learners have preferences that lie -between automatic and on-demand approaches, and draw benefits from each approach depending on certain factors, including their skills in controlling the study sessions, and their familiarity with the subject. The takeaway is that, exploring the design space for hybrid techniques of retrieval practice in video contexts, is a promising way forward.

# Chapter 5

# Exploring the Hybrid Spectrum with HYBREID

In Chapter 4, our findings showed that the absence of visible cues from the video interface disabled some learners to even decide when to stop for an exercise. This was a major drawback of the on-demand design. The automatic nature of the exercises that necessitated working through all qualifying flashcards also removed the freedom the learners to choose the exercises based on their own needs. This was a key drawback of the automatic design.

In this chapter we explore the design of hybrid techniques for retrieval exercises within video contexts such that learners are able to: use automated scaffolding to support their learning, at the same time, stay in control of their study. We designed HYBREID (HYBrid Retrieval Exercises for Informational ViDeos) - a technique for retrieval exercises that combines a non-obtrusive form of automatic prompts and a granular on-demand access to exercises. The goal of the design is to encourage learners to engage with retrieval exercises, in a more subtle way, by allowing them to visually track their own progress and help them assess their understanding of key concepts in the video. To understand the tendencies of learners to interact with retrieval exercises in the hybrid design space, we studied three versions, each one prompting the learners in different ways. Our findings from the final study provide insights on how to avoid interrupting the learner when they are in a flow and extend support to them when they actually need it or explicitly seek it.

## 5.1 Research Approach

Based on our findings in the previous chapter, we learned that while learners like to have control over their studies, they are unable to do so without some amount of scaffolding. Motivated by this finding, we try to answer the research question:

**RQ:** *How can we design in-context automated retrieval exercises for informational videos such that learners still retain control over their study?*

To answer this question, we used a user-centered design process to develop HYBREID (HYBrid Retrieval Exercises for Informational ViDeos), a hybrid technique for retrieval exercises that combines aspects of automated support through the use of subtle reminders and enables learners to control their study by providing visual cues for time stamped exercises and granular access to individual exercises. To gauge the extent of automated prompts that may be desirable, we designed a second version of HYBREID, combining the visual cues for exercises with a timed pause of 10 seconds at topic breaks, instead of reminders. We compared both of these designs by doing an observational lab study with 5 participants. The task for each participant was to learn the main ideas from the given informational videos by making use of the provided retrieval exercises. We randomly showed the participants the videos with only the visual cues in HYBREID - free choice (control condition) or the ones with accompanying automated prompts: HYBREID - with reminders and HYBREID - with short pause (experimental conditions).

## 5.2 Design Goals

The key findings from the study in the previous chapter informed the design for the hybrid version. We retained some aspects from both the designs in the earlier version that had clear benefits. For example, we kept the visual aid for topic transitions, and also the interleaved nature of the exercises. For the next version, we were motivated by the following design goals:

*D*1. *Encouraging timely reflections through automated nudges:*

One of our main findings in the previous study was that when learners were engrossed in watching the video, despite being aware that they could stop at any

point to use flashcards for reflection, they were unable to decide on a suitable moment to pause. Therefore one of our goals was to help learners identify suitable moments to pause and gently nudge them to engage in a reflective exercise, without necessitating them to take action (as opposed to automatic pauses in REMIND that required user action).

*D2. Enabling learners to self regulate their exercises with visual cues:*

One of the major drawbacks of the automatic prompts in our last study was the inherent lack of control a learner experienced during their study. Thus, one of our new design goals was to allow learners to choose which exercise they wanted to see and when they wanted to see them, so they can regulate their own study and customize the reviews to their needs.

## 5.3 Designing the User Interface of HYBREID

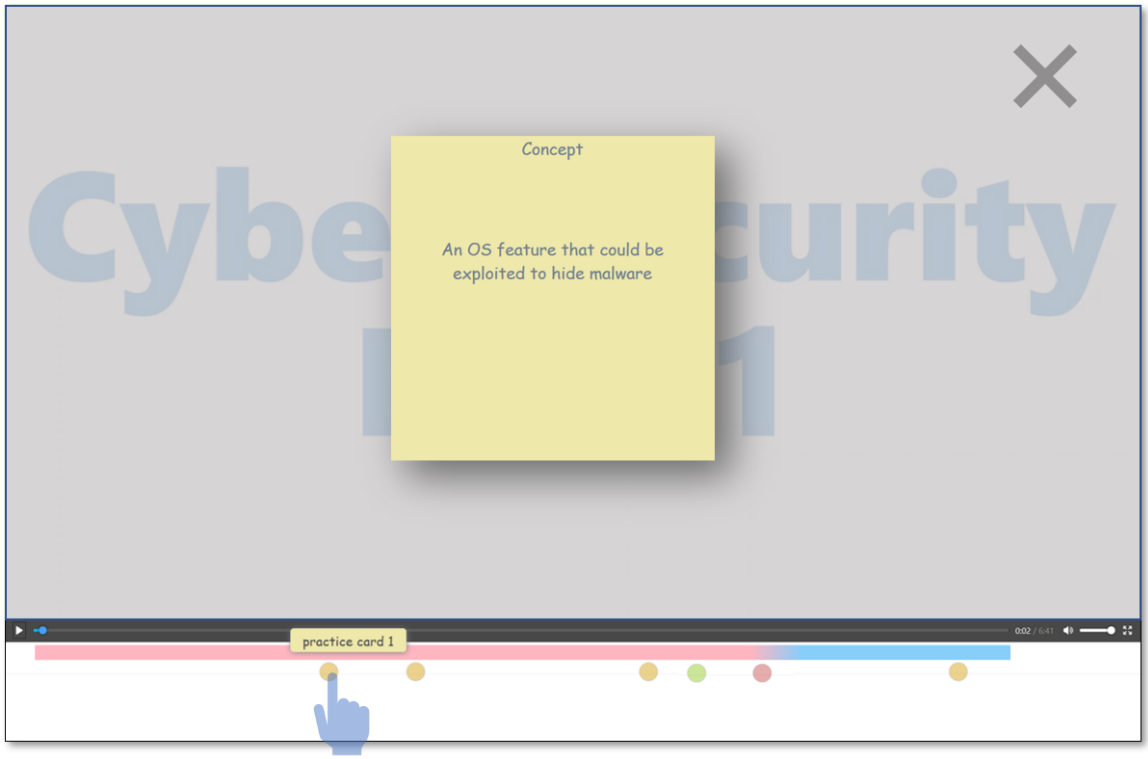


Figure 5.1: The interface of HYBREID showing locations of exercises on the video timeline.

The interface (Fig. 5.1) shows a bar immediately below the video seek-bar. The gradient on this bar indicates a change in topic when the colour transitions. This visual cue on the seek-bar was identified to be useful in helping learners identify a potential moment for pausing and assessing their progress through the subtopics in the video, therefore we decided to maintain this visual cue in the new prototype as well. However, this time we introduced a separate bar for showing this cue, unlike last time where we used the default video progression bar to show the topic transitions - and in doing so, we changed the default appearance of a video interface commonly seen on the web, where the progression stays hidden while the video is in progress. In this prototype, we decided to maintain the default appearance of the video interface, and show the topic transition cues clearly as a separate feature.

We included visual cues for exercises by introducing small dots right below the subtopics bar. Each dot represents a flashcard exercise located at the corresponding time in the video. Each dot also reflect how the learner assessed their progress (Fig. 5.2). If a learner indicates, for any exercise that they "got it", then the corresponding dot turns green, and if the exercise is bookmarked, the dot turns red. If all the cards remain unmarked, they simply remain yellow by default (D2).

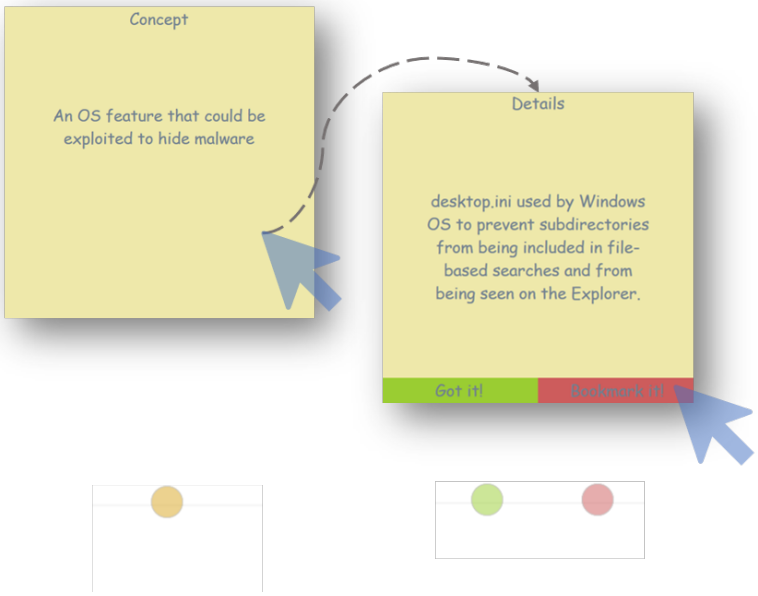


Figure 5.2: A flashcard in HYBREID.

**HYBREID - with reminders:** We added another kind of visual cue, in the form of small reminders that pops up occasionally. Each reminder cue appears for

5 seconds when the play-head passes over the specific timestamp where the exercise is located (Fig. 5.3). This automated reminder is meant to provide a chance for the learner to stop and see exercises if they feel the need for it (D1).


Figure 5.3: HYBREID - with reminders: A reminder cue appears for the learner to see the relevant exercise as the video progresses.

**HYBREID - with short pause:** We also designed a variation of automatic prompting and this design borrows the idea of automatic pauses at topic breaks from the previous study. The difference in this prototype is that in order to maintain the flow of the learner, the pause only lasts 10 seconds and can be dismissed.

For this prototype, we borrowed the contents from the previous study. The cue statements and the explanation statements were the same as the ones used earlier. Except, this time we used shorter videos of 5 to 6 minutes each. We had about 5 - 6 flashcards situated at different locations of the video, located right after the respective concept is discussed.

The system was developed using HTML, Javascript and CSS3. This prototype runs entirely on the client side.

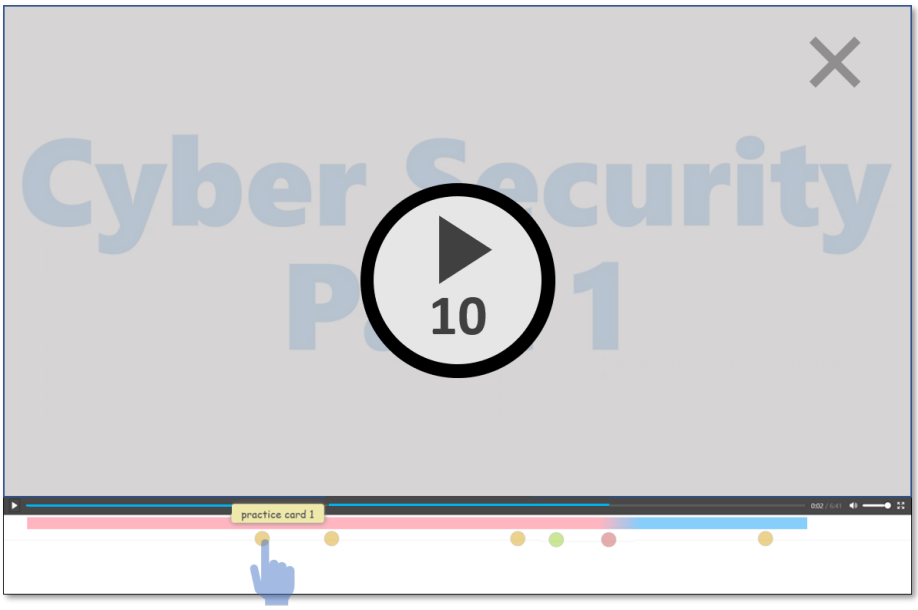


Figure 5.4: HYBREID - with short pause: A countdown of 10 seconds at every topic transition to encourage learners to engage with the exercises

| | | **Test Conditions** |
|---|---|---|
| *Video Chapter* | *Technique* | *Condition* |
| Cyber Security Part 1 | Free Choice | C1 |
| Cyber Security Part 2 | Reminders | C2 |
| Cloud Infrastructure Part 1 | Short Pause | C3 |

Table 5.1: The hybrid techniques with free choice, reminders and short pause were each mapped to different videos

## 5.4 Study Design and Analysis

We used a within-subjects design for our observational study, where each learner used three versions (short-pause, reminders, free-choice) of HYBREID using three different informational videos (Table 5.1). We worked with experts to first identify key concepts in each of the videos and accordingly created a flashcard for each of these concepts (5-6 for each video). Each concept on a flashcard (Fig. 5.2) was associated with a time-stamp to facilitate the retrieval of relevant flashcards at different intervals during playback.

**Participants:** We recruited a new set of 5 participants (2M, 3F), all between the ages of 19-40 who either had a computer science related education, or relevant technical experience. All participants were given a gift card for their participation.

**Study Layout:** Each participant first began the study by filling out a demographic questionnaire. Next, participants interacted with the three different hybrid techniques described above. To counter order effects, all participants were randomly assigned to one of these conditions. We maintained an upper-bound of 1 hour for the entire study. The videos were sourced from Lynda.com (Table 5.1) and were similar to each other in terms of duration, style of presentation and the number of concepts discussed. Each video was nearly 6 minutes long and had 5 to 6 exercises. There was no time limit for the exercise sessions or for completing the surveys. The study was conducted in a lab setting and the Think-Aloud Protocol was followed were the participants were encouraged to share their thoughts during the study.

**Instructions:** Before starting a session, participants were given a tutorial of the prototype. They were also asked not to take notes and were instead encouraged to make the best use of the hints on the flashcard to recall the concept-related details. For the short pause condition, the participants were informed about the 10 second pause where they were free to either stop and take some exercises or dismiss the pause, if they felt they did not need it. For the free-choice and reminders condition, the participants were asked to pause the video on their own and access the flashcards for practice whenever they felt the need for it. We encouraged the participants to use the exercises as they would normally.

## 5.5 Data Analysis

We collected data by directly observing participants' behavior during the study, recording the screens, collecting their feed-back through survey responses and drew further insights from interviews which were recorded and later transcribed. In this evaluation, we allowed learners to choose how many times they wanted to engage with the flashcards exercises. We also made observations to find out whether the learners were cognitively engaging with the exercises by recording their screen and noting their bookmarking behavior, whether they used the short pause to engage with exercises or voluntarily stopped to view an exercise, how often they saw the details in each

card, and also the total time spent in engaging with exercises. Similar to our previous study, we used selected proxy engagement measures [31, 42] which were relevant to our study. We were mainly interested in finding out if the three variations of HYBREID (reminders, short-pause, free choice) were useful, enjoyable, offered a sense of control over their study, boosted confidence about the material, seemed time consuming or distracting and promoted focused attention (Fig. 5.5)

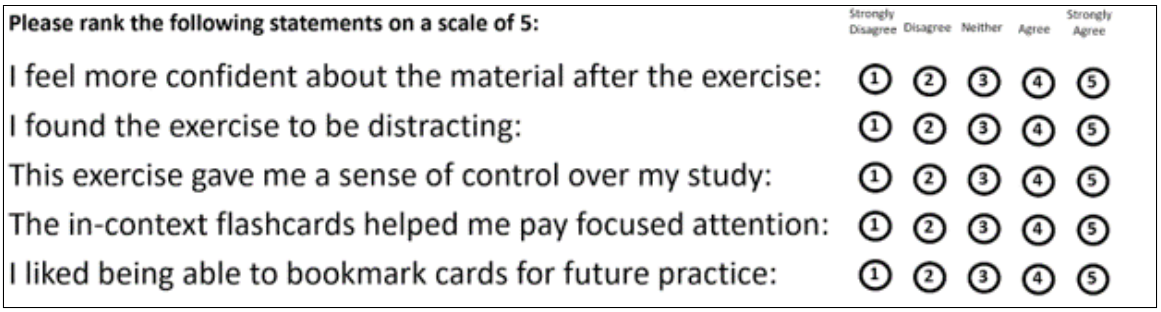


Figure 5.5: The survey questions for evaluating HYBREID.

## 5.6 Key Findings

### 5.6.1 Utility of visual aids for exercises

We observed that the participants were engaged with the exercises in HYBREID. All 5 participants stopped at various points at different points in the videos in all variations of the prototype: at least 4 out of 5 times in HYBREID - with reminders, at least 3 out of 6 times in the HYBREID - with short pause, and at least 3 out of 5 times in HYBREID - with free choice. The behaviour was consistently seen in both the conditions of nudging and even in the absence of it, so it is clear that the presence of visual cues that simply mark the location of relevant exercises enabled the learners to make a decision about whether or not they wanted to stop for an exercise.

> *"It's good that the exercises were spread out because if I don't understand anything I can just look right away"(P05)*

### 5.6.2 Automated prompts help in timely reflections

Most participants (4 out of 5) agreed that both versions of automatic prompting were useful and they displayed taking control of their studies to a great extent in both the experimental conditions (short-pause and reminders) by stopping at various stages of the video to take an exercise. All participants utilized the automatic pause to go back and check the exercises they wanted to see. Some even went back to change a marked card for re-visiting at a later time [P01, P05]. In the control condition where the automatic prompts were completely absent, they relied on the visual cue for the topic change to help them decide on a good moment to stop for exercises.

> *"Here [free choice condition] I had the freedom to stop wherever I wanted so I thought I would stop at the topic change."(P03)*

### 5.6.3 Bookmarking for practice in the next session

One of the broader goals for inserting curated retrieval exercises was to reduce the need for re-watching the videos over and over again, by helping learners reinforce the concepts through practice. While it was not possible to ascertain whether they would re-watch the entire video, we observed that none of the participants went back to the bookmarked the exercises to check them again. On enquiring about this behaviour, we found out that the participants found it more natural to practice the concepts at a later time and also use the bookmarked cards just like one would use an actual bookmark - to guide them to the portion of the video they had not yet grasped.

> *"I bookmarked because I couldn't remember the details and this was for me to revisit later."(P03)*

### 5.6.4 Pros and cons of HYBREID

Participants largely felt that both versions of the automatic prompts (reminders and short-pause) worked for them, for various reasons. One participant sums up how these automatic interventions work well when the choice of engaging with them is with the learners:

> *"[Reminders] tell you where you are in the video and if the things are easy you can just go ahead whereas when things get difficult you know that there is some pause.– if things go faster than you can comprehend then it gives you checkpoints where you can stop"(P03)*

For some participants, the visual cue for relevant exercises did not suffice as they were engrossed in the video and required additional nudging and easier access to the exercises. Reminders were helpful in nudging participants who tend to be lost in the flow of watching the video for good places to stop for an exercise.

> *"when you watch the video, you just watch the video.. you don't remember here's something you need to interact with to get most of the video...stop the video and put a link....because that's more natural.. I think this [reminder] is the closest"(P01)*

All participants unanimously agreed that the reminders did not cause any distractions - this satisfies our broader design goals but on looking closer we discovered that the reminders were not prominent enough for some participants. To some they were "decent sophisticated reminder that was not distracting me" (P02), while some others reported having missed seeing them until they were a long way ahead in the video [P01, P04].

> *"I only saw the reminders at this point [at 3 quarters of the video] that's why I didn't take the exercises before this."(P04)*

One positive finding was that, when the learners were able to exercise some control over their study sessions with the automated cues and curated exercises, they were able to think of more ways to take control of their study. Although 4 out of 5 participants indicated in the survey that they felt they were in control of their study, in the interview they mentioned instances when they felt limited in their learning experience.

> *"I would like things to be customized according to the situation.. timing of the pause, content of the cards."(P02)*

Sometimes participants were able to think of ways of even re-doing the exercises when they saw content that was somewhat similar, but did not exactly align with what they felt they needed to practice at a given time in the video.

> *"I would like to have a flashcard with the 7 steps listed.. I would rather have lists than sentences.. I would have included more summarized version in the details section."(P04)*

### 5.6.5 Areas of Improvement

**User Interaction:** It is worth noticing that, in all the versions we have cases where the participant stopped to check an exercise but did not mark the card. We had designed for the exercises to be marked as understood or to be bookmarked to aid the learners in practicing the concept. Since these exercises were also dismissible, some of them (3 out of 5 participants in each condition) chose simply to see the some of the exercises but mark only a few of the cards they checked. On investigating regarding this behaviour, we realised that it was possible to overlook marking an exercise altogether since they were dismissible, but it was also possible that they were unsure of their own understanding of the content and preferred to leave some cards unmarked within the same study session.

We also observed two participants (in 2 different conditions) change their response from "got it!" to "re-visit". When the topic is new, the learner may have difficulty in gaining sufficient confidence simply from curated in-built exercises [P03, P05].

**Utility of the exercises:** While 3 out of 5 participants agreed that they felt more confident in both cases of automatic prompts (2 out of 5 were neutral in free choice case), One participant P01 was consistently under-confident after the exercises. We investigated the reason and found out that this participant had a very strong notion of how a flashcard exercise should be and the given presentation/content did not suit him: *"In an informational video, it is important to state what the goals are".. (P01)* . P01 was able to frame some examples of what the goals might be by referring to the contents of the flashcard and tailoring it to his understanding.

## 5.7 Summary

In summary, we learned that visual cues of interleaved exercises provide useful support to learners for identifying the key concepts and confirm their understanding of it, while deliberate automated prompts encourage retrieval exercise in learners who find it challenging to engage in practice on their own. The visual cues and automated prompts for curated exercises assist in comprehension of new material, but while granular access to content supports selective practice, without a way to enforce practice, it potentially leads learners to disengage from repeated practice within the same session. Also, when learners have adequate automated support for learning, they are better able to personalize the study session by themselves.

# Chapter 6

# Discussion

This thesis contributes insights into how learners perceive and engage with retrieval exercises in informational videos. Our goal was not to assess the effectiveness of these techniques in promoting learning outcomes, but rather to take a user-centered approach to explore the design of retrieval exercises in video contexts and how they can be made useful. In particular, our findings have shed light on several benefits and drawbacks of designing in-context flashcard-based retrieval exercises and how they should be offered within informational videos. We now reflect on some limitations of our work and highlight several opportunities for future research to explore further innovations and empirical results.

## 6.1 Limitations

One limitation of the studies in this research is that we only assessed retrieval exercises in the context of technical informational videos. Although our interviews with participants suggest that the design of retrieval exercises would be relevant to other subject areas also (like mathematics and history), future work should investigate the relevance of our findings for other types of informational videos. In addition, while all of our participants were adult learners who had experience in learning from online videos, learners in the workplace or other settings where informal learning may take place could exhibit different behaviors and perceptions. We acknowledge that this study used an experimental prototype with basic video features and there are several other ways of designing video-watching interfaces and presenting curated content that

should be further explored. Lastly, in this study, we used proxy measures to gauge engagement. Future work should explore the use of more direct measures for assessing engagement. Nonetheless, our study is a first step towards designing engaging user-centered retrieval-based interventions for enhancing retention in the context of watching informational videos and opens up several promising research directions for further enhancing retrieval-based interventions.

## 6.2 Future work: Supporting automatic in-context retrieval exercises

The design and evaluation of in-context retrieval technique in the first prototype REMIND, showed that overall, learners found the flashcard-based retrieval exercises to be useful and engaging. In particular, the automatically prompted flashcards were most useful as they required less effort and allowed learners to feel more confident about the video content. The automatic in-context format of the exercises helped learners maintain focused attention and prevented them from steering away when they encountered difficulties in comprehension [16].

However, despite the strengths of the automatic approach, a key challenge in designing these exercises is to determine how to appropriately offer them to learners. We observed that although participants rarely missed an automatically prompted exercise, they did like having the choice to dismiss it. One implication here was to design automatic prompts to serve as a gentle reminder at suitable intervals rather than pausing the video. The second prototype HYBREID brought to light that when given the choice of selecting exercises to practice, learners like to eliminate exercises from their practice and this is also consistent with findings in previous research [20]. The exercises they did bookmark, they did so with the intention of re-visiting them in the next study session, rather than in the same one. Our study was set in the laboratory, therefore we only we only did single one-hour session with the participants. In the absence of multiple sessions, we cannot ascertain to what extent the participants would use the bookmarked exercises for practice in the subsequent sessions.

Since earlier studies involving retrieval exercises have established the benefits of spaced out retrieval practice [35], future research could potentially explore variations in spacing these intervals of practice, over multiple study sessions and predicting an

appropriate moment to prompt the learner for taking an exercise over different lengths of duration in informational videos. In addition, future research could also explore how the flashcard content itself could be generated automatically to complement recent efforts in automatic generation of questions in various styles of presentation, such as MCQ [18] and fill in the gaps [6].

## 6.3 Future work: Scaling curating efforts

Another key insight from our findings was that the learners found expert-curated intervals to be helpful in aligning their thoughts rather than figuring out where they should pause-and-reflect entirely on their own. This presents several opportunities to design interventions that provide curated content within the context of a video. For example, future work can consider crowd-sourcing based techniques that have already been leveraged in several studies in HCI in the context of learning [17, 29], such as for aiding video navigation and generating study material for learners [12, 27]. Prior research provides insights into peaks in learner's viewership in the time-line of video [22], which could be used to aid potential contributors [49], such as video-authors or domain-experts, to locate suitable points for inserting exercises. Potentially, peers can also contribute content which could then be moderated by experts. There is already an active community of domain experts who engage in contributing content online, such as in the forums. Currently, in-context exercises are integrated within video only before production. However, using automated web technologies there's a chance to augment the videos with simple reflective exercises based on crowd-sourced locations where learners actually could benefit from reflection. Future work could look into ways of tapping into the potential of the learner community to aid them to extend their ways of contributing content for informational videos.

## 6.4 Future work: Personalizing exercises with hybrid designs

Our study with REMIND revealed that learners like to be able to attempt an exercise when they start feeling overwhelmed or find their attention waning. Several learners also pointed out that they like their study to be more targeted and wanted to be able

to choose which of the curated exercises to attempt or practice. Although they liked to be able to bookmark a flashcard to revisit it later, they felt it was unnecessary to practice other items every time.

Learners also described pain-points with a concept that they may experience in between two automatically curated exercises, highlighting one of the strengths of the control offered by the on-demand design. These findings point toward a retrieval exercise technique that is a hybrid between an automatically guided routine and an on-demand approach. The automatic approach offers learners the chance to refine their understanding of the material, but a hybrid design, with on-demand exercises, could enable learners to make their study more selective and directed. Pain-points may differ among individuals and a hybrid of these two designs provides a chance to personalize their reflection and self-assessment. This behaviour was even more clearly observed during their interaction with HYBREID, which was designed precisely to support learners make selective practice choices. Future research should explore ways to encourage learners to consciously engage in more practice, as the benefit of retrieval exercise improves with more practice.

Future research should not only look at ways to personalize the retrieval exercises automatically, but also direct efforts towards enabling learners to customize the default content. As we observed in our studies with both REMIND and HYBREID, automated interventions provide the initial scaffolding that some learners require in order to comprehend and engage actively with video content, but once they are able to do so, they often are able to think of ways to fine-tune the exercises to their own specific needs. From this study, we have some insights on how retrieval exercises might boost engagement with informational videos across learners with different abilities for controlling their study. More research needs to be undertaken to understand how learners might take control of the personalizations themselves. For example, a direction for personalizing retrieval exercises in the future could be to design approaches for learners to create their own in-context flashcards, or to customize expert curated flashcards to facilitate and complement their own note-taking processes.

## 6.5 Future work: Supporting meta-cognitive skills for boosting practice

Previous research in educational psychology shows that people with higher meta-cognitive skills are able to better control their study sessions and are therefore more engaged with the material and those with lower skills could be supported with some scaffolding efforts [23]. We have similar findings in our study with REMIND and HYBREID as well, where we observed that while certain learners look out for opportunities to practice, some others are not conscious about practicing even though they are aware of the exercises and the benefits of taking them. Prior research in video-based learning contexts [46] have also suggested taking efforts towards assisting learners in developing some meta-cognitive skills that will enable them to take conscious decisions to practice. Since the benefits of retrieval practice can be gained only when there is some repeated self-testing, and automating the prompts is one way of encouraging that, the other possible direction is in exploring ways to enable learners to identify their own goals, before identifying their pain-points. This could simply be done by presenting the outline of a video. In our studies participants reported experiencing some difficulty orienting their expectations from the video without a clear outline of the content. Future research could explore ways helping learners make progress towards their goals by increasing opportunities for retrieval practice.

# Chapter 7

# Conclusion

In this thesis, we have contributed insights into how learners perceive and engage with retrieval exercises in informational videos. Our formative study provides initial design goals leading to the development of the first prototype REMIND, that offers curated interactive flashcards within the context of a video. This technique allows learners to stop the video at suitable moments and practice retrieval exercises to reinforce their learning. The comparative study we ran using automated and on-demand versions of this prototype offers insights into the pros and cons of the two techniques, and contributes implications for exploring hybrid designs for retrieval exercises. The design implications from the study with REMIND, led us to develop HYBREID, a second prototype for flashcard based exercises for informational videos, that borrows the favorable aspects of automated prompts and on-demand interactions of REMIND. The second study with HYBREID strengthens our earlier findings and provides further implications for supporting hybrid techniques for designing retrieval exercises, pointing us in the direction of providing automated support when it is needed, while enabling learners to ultimately take control of their own study session. Researchers and designers interested in pushing the boundaries of learning experiences online, will find potential directions for future work, like: to what extent to automate learning support, how to help learners take control of their study, and how to leverage the resources available in the community to support learning in informal contexts.

# Bibliography


[1] Retrieval practice and study planning in moocs: Exploring classroom-based self-regulated learning strategies at scale. pages 57–71, 2016.

[2] Coursera, 2019.

[3] Edx, 2019.

[4] Edxassessments, 2019.

[5] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. Engaging with massive online courses. In *Proc. international conference on World wide web*, pages 687–698. ACM, 2014.

[6] Lee Becker, Sumit Basu, and Lucy Vanderwende. Mind the gap: learning to choose gaps for question generation. In *Proc. Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 742–751, 2012.

[7] Janell R. Blunt and Jeffrey D. Karpicke. Learning with retrieval-based concept mapping. volume 106, pages 849–858, 2014.

[8] Juliet M Corbin and Anselm Strauss. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, 13(1):3–21, 1990.

[9] William L Cull. Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 14(3):215–235, 2000.

[10] Dan Davis and Claudia Hauff. The Half-Life of MOOC Knowledge A Randomized Trial Evaluating Knowledge Retention and Retrieval Practice in MOOCs. *Proc. Learning Analytics and Knowledge*, pages 1–10, 2018.

[11] Degreed.com. How the workforce learns in 2016, 2016.

[12] Vania Dimitrova, Antonija Mitrovic, Alicja Piotrkowicz, Lydia Lau, and Amali Weerasinghe. Using learning analytics to devise interactive personalised nudges

for active video watching. In *Proc. User Modeling, Adaptation and Personalization*, pages 22–31. ACM, 2017.

[13] Samuel Dodson, Ido Roll, Matthew Fong, Dongwook Yoon, Negar M Harandi, and Sidney Fels. An active viewing framework for video-based learning. In *Proc. L@S*, page Article 24. ACM, 2018.

[14] Darren Edge, Stephen Fitchett, Michael Whitney, and James Landay. Memreflex: Adaptive flashcards for mobile microlearning. In *Proc. MobileHCI '12*, pages 193–194. ACM, 2012.

[15] Luke G. Eglington and Sean H.K. Kang. Retrieval Practice Benefits Deductive Inference. *Educational Psychology Review*, 30(1):215–228, 2018.

[16] John Frens, Erin Walker, and Gary Hsieh. Supporting answerers with feedback in social q&a. In *Proc. L@S*, page Article 10. ACM, 2018.

[17] Neil T Heffernan, Korinn S Ostrow, Kim Kelly, Douglas Selent, Eric G Van Inwegen, Xiaolu Xiong, and Joseph Jay Williams. The future of adaptive learning: Does the crowd hold the key? *International Journal of Artificial Intelligence in Education*, pages 615–644, 2016.

[18] Yi-Ting Huang, Ya-Min Tseng, Yeali S Sun, and Meng Chang Chen. Tedquiz: automatic quiz generation for ted talks video clips to assess listening comprehension. In *IEEE Conference on Advanced Learning Technologies (ICALT)*, pages 350–354, 2014.

[19] Sean HK Kang, Tamar H Gollan, and Harold Pashler. Don't just repeat after me: Retrieval practice is better than imitation for foreign vocabulary learning. *Psychonomic bulletin & review*, 20(6):1259–1265, 2013.

[20] Jeffrey D Karpicke. Metacognitive control and strategy selection: deciding to practice retrieval during learning. *Journal of Experimental Psychology: General*, pages 469–486, 2009.

[21] Juho Kim, Elena L Glassman, Andrés Monroy-hernández Meredith, and Ringel Morris. RIMES : Embedding Interactive Multimedia Exercises in Lecture Videos. *CHI*, pages 1535–1544, 2015.

[22] Juho Kim, Philip J. Guo, Daniel T. Seaton, Piotr Mitros, Krzysztof Z. Gajos, and Robert C. Miller. Understanding in-video dropouts and interaction peaks in online lecture videos. *Proc. L@S*, pages 31–40, 2014.

[23] René F Kizilcec, Mar Pérez-Sanagustín, and Jorge J Maldonado. Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers & education*, 104:18–33, 2017.

[24] René F Kizilcec, Chris Piech, and Emily Schneider. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the third international conference on learning analytics and knowledge*, pages 170–179. ACM, 2013.

[25] Nate Kornell and Kalif E Vaughn. How retrieval attempts affect learning: A review and synthesis. In *Psychology of learning and motivation*, volume 65, pages 183–215. Elsevier, 2016.

[26] Geza Kovacs. QuizCram: A Question-Driven Video Studying Interface. *Extended Abstracts CHI'15*, 2:133–138, 2015.

[27] Ching Liu, Juho Kim, and Hao-chuan Wang. ConceptScape : Collaborative Concept Mapping for Video Learning. *CHI*, pages 1–12, 2018.

[28] Mark A McDaniel, Pooja K Agarwal, Barbie J Huelser, Kathleen B McDermott, and Henry L Roediger III. Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology*, pages 399–414, 2011.

[29] Piotr Mitros. Learnersourcing of complex assessments. In *Proc. L@S*, pages 317–320. ACM, 2015.

[30] Toni-Jan Keith Palma Monserrat, Yawen Li, Shengdong Zhao, and Xiang Cao. L.IVE: an integrated interactive video-based learning environment. *Proc. CHI*, pages 3399–3402, 2014.

[31] Heather L O'Brien, Paul Cairns, and Mark Hall. A practical approach to measuring user engagement with the refined user engagement scale (ues) and new ues short form. *International Journal of Human-Computer Studies*, 112:28–39, 2018.

[32] Yeonjeong Park, Insung Jung, and Thomas C Reeves. Learning from moocs: A qualitative case study from the learners' perspectives. *Educational Media International*, 52(2):72–87, 2015.

[33] Bernhard Pastötter and Karl Heinz T. Bäuml. Retrieval practice enhances new learning: The forward effect of testing. *Frontiers in Psychology*, pages 1–5, 2014.

[34] Mary A Pyc and Katherine A Rawson. Examining the efficiency of schedules of distributed retrieval practice. *Memory & Cognition*, 35(8):1917–1927, 2007.

[35] Henry L. Roediger and Andrew C. Butler. The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1):20–27, 2011.

[36] Henry L Roediger III and Jeffrey D Karpicke. The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, pages 181–210, 2006.

[37] Henry L Roediger III and Jeffrey D Karpicke. Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological science*, pages 249–255, 2006.

[38] Ralf Schmidmaier, Rene Ebersbach, Miriam Schiller, Inga Hege, Matthias Holzer, and Martin R Fischer. Using electronic flashcards to promote learning in medical students: retesting versus restudying. *Medical education*, pages 1101–1110, 2011.

[39] Niels Seidel. Video Assessment Patterns: Designing Video Player for integrated Assessment. *EuroPlop '16: Proc.*, pages 1–20, 2016.

[40] Burr Settles and Brendan Meeder. A Trainable Spaced Repetition Model for Language Learning. *Proc. Association for Computational Linguistics*, pages 1848–1858, 2016.

[41] Sara Shrader, Maryalice Wu, Dawn Owens, and Kathleen Santa Ana. Massive open online courses (moocs): Participant activity, demographics, and satisfaction. *Online Learning*, 20(2):199–216, 2016.

[42] Vivek Singh, Balaji Padmanabhan, Triparna de Vreede, Gert-Jan de Vreede, Stephanie Andel, Paul E Spector, Steve Benfield, and Ahmad Aslami. A content engagement score for online learning platforms. In *Proc. L@S*. ACM, 2018.

[43] Megan A. Smith and Jeffrey D. Karpicke. Retrieval practice with short-answer, multiple-choice, and hybrid tests. *Memory*, 22(7):784âĂŞ802, 2013.

[44] Megan A Smith and Jeffrey D Karpicke. Retrieval practice with short-answer, multiple-choice, and hybrid tests. *Memory*, pages 784–802, 2014.

[45] Tim Van Der Zee, Dan Davis, Nadira Saab, Bas Giesbers, Jasper Ginn, Frans Van Der Sluis, Fred Paas, and Wilfried Admiraal. Evaluating Retrieval Practice in a MOOC: How writing and reading summaries of videos affects student learning. *Proc. Learning Analytics and Knowledge*, (Article 10):216–225, 2018.

[46] Ourania Maria Ventista. Self-assessment in massive open online courses. *E-Learning and Digital Media*, 15(4):165–175, 2018.

[47] Sarah Weir, Juho Kim, Krzysztof Z Gajos, and Robert C Miller. Learnersourcing subgoal labels for how-to videos. In *Proc. CSCW*, pages 405–416. ACM, 2015.

[48] Joseph Jay Williams and Juho Kim. Understanding the Effect of In-Video Prompting on Learners and Instructors. *Proc. CHI*, pages 1–12, 2018.

[49] Joseph Jay Williams, Juho Kim, Anna Rafferty, Samuel Maldonado, Krzysztof Z Gajos, Walter S Lasecki, and Neil Heffernan. Axis: Generating explanations at scale with learnersourcing and machine learning. In *Proc. L@S*, pages 379–388. ACM, 2016.

[50] Kathryn T Wissman, Katherine A Rawson, and Mary A Pyc. How and when do students use flashcards? *Memory*, 20(Article 6):568–579, 2012.